DOCUMENT RESUME

ED 156 215 IR 006 448

AUTHOR Griswold, George H.; Kapp, George H.
TITLE Computer Processed Evaluation.
PUB DATE Jun 77
NOTE 12p.; Paper presented at the Conference on Computers in the Undergraduate Curricula (East Lansing, Michigan, June 20-22, 1977); For entire proceedings, see IR 006 142

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Chemistry Instruction; *Computer Assisted Instruction; Cost Effectiveness; Educational Innovation; Electronic Data Processing; Evaluation Methods; *Mastery Tests; *Student Testing; Systems Approach; *Test Construction; Test Items; Undergraduate Study
IDENTIFIERS Washtenaw Community College MI

ABSTRACT

A student testing system was developed consisting of computer generated and scored equivalent but unique repeatable tests based on performance objectives for undergraduate chemistry classes. The evaluation part of the computer system, made up of four separate programs written in FORTRAN IV, generates tests containing varying numbers of multiple format questions and provides students with immediate feedback on a cost effective basis. Appendices contain examples of a computer generated test, a computer scored test, and the computer generated distribution of total test performances for both an individual and a class. (Author/CHV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED156215

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

COMPUTER PROCESSED EVALUATION

BY

George H. Griswold

and

George H. Kapp

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ted Sjoersma

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR006448

Computer Processed Evaluation

George H. Griswold[1] and George H. Kapp
Chemistry Department
Washtenaw Community College
Ann Arbor, Michigan 48106
(313) 971-6300

Abstract

This paper deals with a system of computer generated and scored equivalent, but unique repeatable tests based on performance objectives, containing varying numbers of multiple format questions which provide for immediate feedback to students on a cost effective basis, which was developed in the Department of Chemistry at Washtenaw Community College, Ann Arbor, Michigan.

Introduction

Faced with a constantly expanding student population and all of the problems which are associated with it, the Department of Chemistry at Washtenaw Community College has turned to the development and use of a systems approach utilizing the computer for those tasks which are objective and clerical in nature. The system identified four key parts: (1) objectives, to formally set the goals to be achieved, (2) the instructor, to interpret the objectives to the student, (3) the student to receive the instruction, and (4) evaluation to provide the feedback for the other parts of the system. The primary use of the computer is in the evaluation part of the system where it releases the instructor from the clerical routine of generating and scoring test instruments, and permits him to concentrate on the instructional part of the educational process.

The cornerstones of our system are the sets of performance objectives, thus the other parts of the system are keyed to them. At the beginning of each course, the students are informed that the objectives for each unit of study will be duplicated and distributed. The students are then encouraged to use the objectives in a "checkoff" manner as each unit progresses, as well as in their final review prior to taking the test on that unit. Further, it is made clear to the students that all of the questions will be drawn directly from the objectives. With the students keyed in on the objectives, the instructor and the evaluation systems must also be keyed to the objectives.

System Components

The evaluation part of our system consists of four separate programs written in FORTRAN IV: (1) a LOADER which reformats the test questions, (2) a TEST GENERATOR which creates tests according to a given set of specifications, (3) a GRADER which evaluates student responses and records student grades, and (4) a GRADEBOOK for editing and summarizing student records. Using these four programs on the University of Michigan Terminal System (MTS) we have been able to implement the evaluation part of our system for two courses—Introductory Chemistry 057, and General Chemistry III.

In addition to the four operating programs, our system includes the questions which are associated with the seventeen units divided between the two implemented courses. Our test questions are each keyed by objective and unit, and maintained in machine readable form with each of the seventeen units being held in a separate file. The questions were originally entered into disk files using a terminal to take advantage of the special and lower case characters which are available on the IMB TN print chain which was to be used for output of the tests. The number of questions in each unit's file (bank) varies with the

number of objectives in the unit, and the number of questions (versions) attached to each objective. While the number of versions varies from one to nine—the maximum permitted—most objectives currently in our banks have five questions attached. Because of the large volume of questions involved in the seventeen units, they have been copied from their original location on disk files to magnetic tape for back-up and storage purposes. While the question banks are on disk file, they are usually edited using the MTS EDITOR. In this way the banks are maintained in a current and "living" condition, which we feel has been vital to the success of our system.

Test Generation

The way in which our TEST GENERATOR program selects questions for tests is based on a different philosophy than other systems which were reported at the last CATC Conference. In establishing a set of performance objectives and distributing them to the students all of those things which the student will be expected to be able to do have been established. Thus, an evaluation of those items should in theory be attempted on a test. However, in dealing with a classroom situation where time is a limiting factor, it is not possible to test each student on all of the objectives. Therefore, rather than make a subjective decision as to which objectives are "more important" than others, our TEST GENERATOR has been programmed to select questions for a given test form by first randomly selecting an objective, and then randomly a question attached to that objective (version), without selecting the same objective twice for any one test form. By having the test questions selected in this manner, two things are accomplished: (1) The student is forced to study all of the objectives which promotes more learning. (2) The probability of student complaints as to the content which was covered by the test are greatly reduced. Using our system the only subjective decision the instructor need make is whether a particular question should be placed in the bank and attached to a specific objective.

Another unique feature of our TEST GENERATOR is its ability to accept questions of more than one type. Formally, the TEST GENERATOR is programmed to accept four types of questions: (1) Fixed answer, (2) multiple choice, (3) random number, and (4) double random number. The fixed answer type of question can be used for: true-false, multiple true-false, multiple choice, matching, fill-in, and key word formats when the proper wording and instructions are used with the question. This type of question is limited at the present time to six responses, but could easily be expanded to nine.

Figure 1

A Fixed Answer Question

Select "O" for oxidation, or "R" for reduction, as the name of the process described below:

A loss of electrons.

The multiple choice type of question is used for questions having one answer and four distractors of not more than thirty characters each. As the test forms are generated, the answer and distractors are rotated randomly among the referencing letters (A) through (E), while the identity of the answer's position is maintained for grading purposes. When either the answer of the distractors require more then thirty characters, multiple versions of the fixed answer type can be used to accomplish the same result.

Figure 2

A Multiple Choice Question

Select one of the following as the oxidation number for the ion shown below:

Na^{1+}

(A) +2
(B) 0
(C) +1
(D) +3
(E) −1

The random number type is used for calculation problems whenever multiple versions of the same type of calculation are

desired. The problem is worded in a way as to reference the random number which is printed below the problem according to a given set of specifications including: form (decimal, or scientific), value, number of digits, and unit (limited to one line of thirty characters). The answer is calculated using a given set of specifications including: the mathematical equivalent of all of the other values in the problem, and whether to multiply, or divide this value into the random number, saving the answer for grading purposes.

Figure 3

A Random Number Question

Calculate the number of grams in the number of milligrams shown below:

Report the first digit to the right of the decimal point in your answer.

6812.1 milligrams

The double random number type is used for questions involving the simple mathematical operations of addition, subtraction, multiplication (also squaring), and division of two numbers. The problem is worded in such a way as to reference the two random numbers which will be printed, each enclosed in parenthesis with a specified mathematical sign between them, below the question, according to specifications including: form (decimal, or scientific), value, number of digits, and with the first number equal to, less than, or greater than the second number. The answer to the problem is calculated using the specified operation, and maintained for grading purposes.

Figure 4

A Double Random Number Question

Calculate the answer to the problem below: Rewrite your answer in standard exponential form if necessary, and then report, in order:

a. The first digit to the left of the decimal point in your answer.

b. The sign of your exponent, using "P" for +, or "N" for −.

c. The exponent of your answer.

$(4.5743 \times 10^{+06}) + 2.5297 \times 10^{+03})$

As we were interested in producing repeatable tests, our TEST GENERATOR was programmed to create test forms of equal difficulty, but not necessarily equal numbers of questions. By generating test forms of equal total difficulty, but with varying numbers of questions we have been able to almost completely eliminate any security problem created by the release of old test forms. Using this concept any one test form may contain a large number of "easy" questions, a small number of "difficult" questions, or some random mix of "easy" and "difficult" questions. To obtain the proper number of questions of the right difficulty for each test form, each question is assigned a number of points subjectively to reflect its difficulty. One point is assigned to those questions which are considered to be "easy." Two, three, or more points may be assigned to more difficult questions. the assignment of additional points to a question usually anticipates that it will take an average student two, three or more times longer to consider, calculate, or otherwise develop a correct response to that question. With each available question having been assigned a number of difficulty points, the TEST GENERATOR can select questions randomly according to objective and version until a specified number of points is reached to complete each test form, thus creating unique tests with equivalent difficulty which are repeatable and present a minimum security problem.

The actual process of test generation using our system is extremely simple—after the questions have been written, assigned to an objective, given a number of difficulty points, and placed in a machine readable form on disk file. Using the CONTROL feature of the MTS system which premits batch jobs to be queued for off hours processing from a terminal, we can initiate the generation of any number of completely unique test forms in a terminal session which lasts no more than five minutes. To generate the tests the LOADER program must be run to reformat the question bank. Normally the only output produced at the terminal by the LOADER is a state-

ment of the total number of objectives in the bank being used. (This information is useful as the TEST GENERATOR was programmed to limit the size of each test form to a number of questions which is 80 percent of the number of objectives in the bank.) However, if problems in reformating the bank are detected specific error messages will be displayed. Once the bank has been reformatted, the TEST GENERATOR can be run, and the four specifications necessary entered in order: (1) the number of tests to be generated, (2) the total number of points for each test form, (3) the first line of the test title, and (4) the second line of the test title. Following the entry of these specifications there is little to do until the test forms are ready to be picked up from the line printer.

The printed test forms themselves contain several notable features. The most notable is the printing of the answer(s) to each question on the right-hand side of the paper. A heading, is printed at the beginning of each test form on both the test and answer sides of the paper containing: the test form number, and the test title. To aid in the separation of the answers from the test form, a cut mark is printed in each line of printing. When the answers are separated from the test forms, the test forms take on the usual 8½ by 11 format. The format used for each question includes the printing of the question number, number of points assigned to the question, and the response number(s) (position on the answer sheet) for the question. Following the basic information, the instructions, text, and any required random number(s) are printed. Another notable feature of the printed test forms is the separation of questions between pages. (Unless a question exceeds one page in length—the LOADER program will accept questions with up to ninety-nine lines of text.) If the printing of the next question can not be completed on the current page, the rest of that page is skipped and the printing is continued at the top of the next page. To aid in the separation of one test form from another, a line of printing is placed on the crease between the bottom of one test form and the top of the next test form. Through the use of these notable features it is possible to print test forms and answers which contain all the essential information and are yet readily separable into tests and answers.

Use and Scoring of Tests

In our system the tests are administered during a regularly scheduled class period of fifty-five minutes. The students are expected to have available at these sessions a hand-held electronic calculator and a pencil for marking the answer sheet. As each test form is unique, the test forms are given out one at a time with an answer sheet to prevent those students having the belief that one test form may be easier than another from sorting through the stack and holding up the distribution process. The students are encouraged to set-up any calculations on their test form and then use their calculator to obtain the answers, writing them down directly under where the response number(s) are shown on the test form. Only after all of the questions have been answered in this manner, is the student encouraged to fill out the answer sheet. When the student has completed and checked the answer sheet, it is checked again by the instructor to verify that the student has recorded an identification number, test form number, and that the number of answers given meets or exceeds the number of points assigned to that test. Immediately, after checking the answer sheet, the student is handed the answers to the test form indicated on the answer sheet. (The student always retains the test form.) Through this procedure, the student is immediately able to score the test by comparing the answers with those responses previously recorded on the test form, and determine the percentage score which will be recorded later when the answer sheets are processed. Currently we offer each student two additional opportunities to take a test on the same unit of study during prescheduled periods outside of the regular class meeting times. This procedure of test administration eliminates most of the need to evaluate excuses for absenteeism on test days, and has the added benefit

of providing a mechanism whereby the tests can be used as a learning tool.

In addition to the generation of tests, our system also scores and records the grades obtained by the students. As the TEST GENERATOR creates the test forms, it stores the answers in a disk file. The scoring process starts with the conversion of the answer sheets into machine readable form (currently punched cards from an optical scanner). The GRADER reformats the answers on the disk file and then compares them sequentially with the student responses on a match or mismatch basis. One point is given for a match, and no points for a mismatch. The points earned are then divided by the total points possible and a percentage calculated. This percentage is then recorded in the disk file where the student records are kept. With the repeatable feature of the test forms, we have programmed the GRADER to retain only the highest percentage obtained on each of the test forms of a given test taken by a given student. Thus under this system the student has nothing to lose by repeating a test. After the grade has been recorded, a grade report is generated which gives the test form number, the student's identification number, test title, question numbers, response numbers, the student's responses, and the answers. In addition the report shows the points earned, the total number of points possible, and the percentage. This report is distributed to the student so that he/she can verify the questions which he/she missed, and the grade which was obtained. The GRADER also produces a summary report for posting. The summary report shows the test title, the date on which the tests were scored, the average percentage grade for all the tests scored, and a listing of the identification number of each individual and corresponding percentage grade, sorted by identification number. With this report, the students can determine how they did on all of the tests they took, whether they improved or not, and how their tests compared to all of the tests which were scored. The GRADER also produces a third type of report which shows the total number of responses which were expected for each question in the bank for that test, and the number of correct responses which were received. This report is useful in evaluating how randomly the TEST GENERATOR is selecting questions. This report can also be used to evaluate whether a question is too difficult, faulty in some other manner, or whether the instructor failed to present the material in an effective manner. Through the use of the GRADER we are able to score and record the grades for a large number of unique tests, and produce three different kinds of reports which permit further evaluation, and feedback into our system.

In special cases we do hand-grade a few tests using the answer side of the printed form before giving it to the student. The problem of recording the grades from these hand-graded tests prompted the development of our GRADEBOOK program which gives us the capability to edit the student record file. With the GRADEBOOK we can change a student's identification number, name, or section number as well as any of the grades obtained by that student. Using the GRADEBOOK we can also add or delete complete student records to the student record file. An additional feature of the GRADEBOOK is its ability to permit a complete set of grades to be added to all of the records in the student record file when a book report or other non-test grade is part of the course. However, the most useful feature of the GRADEBOOK is its ability to compute weighted percentages using the recorded grades for each student at the end of a marking period. Through the use of the GRADEBOOK, an instructor can modify a student's grades to reflect any special circumstances which may occur, and also compute accurately and rapidly course grades at the end of a marking period according to the instructor's subjective evaluation as to how each grade should be weighted relative to the total course grade.

Economics and Summary

This past fall semester we generated over 3,000 tests having over 50,000 questions for

the students in the two courses. With three opportunities to take each test, we found that it was necessary to generate a number of tests which was 1.5 times the number of students active in the course. Since we experienced a change in the rate structure at MTS during the latter part of the semester (a more favorable rate was implemented for those willing to process later at night), it is difficult to give any exact cost figures for the entire semester. However, under the new rates, our cost to generate each question appears to be on the order of .6 cents depending somewhat on the number of text lines which have to be printed. Our cost of scoring, and recording of the tests is on the order of .1 cents per question when the tests are processed in batches of 200 or more, but does not exceed .2 cents per question when as few as ten tests are processed. As our answer sheets permit no more than twenty-eight responses, the expense involved in generating and scoring a test of twenty-five responses (the maximum we use) is on the order of 25 cents when all of the additional costs have been figured. Based on these figures and the educational gains which are possible using this system, we feel that it has proven to be cost effective.

Note

1. All correspondence with regard to the contents of this paper should be directed to Professor Griswold.

FORM NUMBER 2

CHEMISTRY 57
TEST NO. 3

NAME _____

1. (3 POINTS)

Calculate the answer to the problem below:
Rewrite your answer in standard exponential form if necessary, and then report, in order:
a. The first digit to the left of the decimal point in your answer.
b. The sign of your exponent, using "P" for +, or "M" for - .
c. The exponent of your answer.

(9.9205 X 10+00) + (8.3701 X 10+03)

2. (1 POINT)

Convert the following number to decimal form.
Report the first digit to the left of the decimal point in your answer.

3298.9 X 10+00

3. (1 POINT)

Select one of the following as the way in which 2.56 should be written in standard exponential form.

(A) 2.56 X 10+0
(B) none of them
(C) 256.0
(D) 2.56 X 1
(E) 2.56

4. (1 POINT)

Write the following number in standard exponential form.
Report the first digit to the right of the decimal in your answer.

209.18

RESPONSES(1- 3)

1.
9.92885E+06

RESPONSE(4)

2.
3298.908

RESPONSE(5)

3.
(A) 2.56 X 10+0

RESPONSE(6)

4.
2.09177E+02

FORM NO. 2

CHEMISTRY 57
TEST NO. 3

FORM NO. 57

CHEMISTRY C57
TEST NO. 8

WARD CHERYL D
JOURNAL NO. 404

QUESTION NUMBER	RESPONSE NUMBER	REPORTED ANSWER	CORRECT ANSWER	POINTS
1	1	A	A	1
2	2	4	4	1
3	3	T	T	1
4	4	3	1	**
	5	3	3	1
	6	2	2	1
5	7	B	B	1
6	8	2	2	1
	9	8	8	1
7	10	T	T	1
8	11	F	F	1
9	12	3	3	1
10	13	3	5	**
	14	1	4	**
	15	C	6	**

TOTALS POINTS = 11

POINTS POSSIBLE = 15

PERCENTAGE = 73

APPENDIX B

GRADE REPORT

CHEMISTRY III TEST NO. 5 DATE DEC 10, 1976

NUMBER OF TESTS GRADED = 71

AVERAGE PERCENTAGE FOR ALL TESTS = 73.500

JOURNAL #	PERCENT	JOURNAL #	PERCENT	JOURNAL #	PERCENT	JOURNAL #	PERCENT	JOURNAL #	PERCENT
3	64	35	84	83	84	111	68		
3	48	37	72	83	84	111	56		
5	84	39	72	84	84	11[illegible]	48		
9	96	39	96	84	93	116	64		
9	92	44	81	86	88	116	100		
12	81	44	92	86	88	117	88		
16	80	49	72	89	52	118	62		
18	81	50	56	90	38	118	88		
18	76	53	68	90	48	119	72		
19	48	53	96	92	44	119	84		
19	60	55	60	92	58	448	89		
20	92	58	96	93	84				
21	60	60	80	94	40				
21	88	66	76	94	68				
23	84	69	80	101	96				
24	76	74	68	103	56				
24	73	80	38	103	84				
29	68	80	72	107	64				
29	72	82	56	110	80				
33	48	82	60	110	85				

11

RESPONSE DISTRIBUTION CHEMISTRY 057

TEST NO. 1 DATE SEP 23, 1976

OBJ	VER	TOTAL	RIGHT
1	1	15	7
1	2	9	6
1	3	8	6
1	4	9	8
1	5	10	3
1	6	5	5

OBJ	VER	TOTAL	RIGHT
2	1	17	12
2	2	22	20
2	3	15	13

OBJ	VER	TOTAL	RIGHT
3	1	28	24
3	2	22	19

OBJ	VER	TOTAL	RIGHT
4	1	15	15
4	2	21	19
4	3	10	9

OBJ	VER	TOTAL	RIGHT
5	1	28	17
5	2	30	19

OBJ	VER	TOTAL	RIGHT
6	1	25	18
6	2	25	20